

# **A forward modelling approach for interpreting impeller flow logs**

*Alison H. Parker<sup>1</sup>, L. Jared West<sup>2</sup>, Noelle E. Odling<sup>2</sup>, Richard T. Bown<sup>3</sup>*

*1. Centre for Water Science, Building 39, Cranfield University, Cranfield, MK43 0AL, UK, a.parker@cranfield.ac.uk*

*2. School of Earth and Environment, University of Leeds, Woodhouse Lane, Leeds, LS2 9JT, UK*

*3. Innovia Technology, St Andrew's House, St Andrew's Road, Cambridge, CB4 1DL, UK*

## **Abstract**

A rigorous and practical approach for interpretation of impeller flow log data to determine vertical variations in hydraulic conductivity is presented and applied to two well logs from a Chalk aquifer in England. Impeller flow logging involves measuring vertical flow speed in a pumped well, and using changes in flow with depth to infer the locations and magnitudes of inflows into the well. However, the measured flow logs are typically noisy, which leads to spurious hydraulic conductivity values where simplistic interpretation approaches are applied. Here a new method for interpretation is presented which first defines a series of physical models for hydraulic conductivity variation with depth and then fits the models to the data, using a regression technique. Some of the models will be rejected as they are physically unrealistic. The best model is then selected from the remaining models using a maximum likelihood approach. This balances model complexity against fit, for example using Akaike's Information Criterion.

24

## 25 **Introduction**

26 Characterizing hydraulic conductivity and its distribution is a key requirement in the  
27 development of groundwater flow models used by regulatory bodies and water companies  
28 to predict pollutant travel times and to design extraction wells, ground water monitoring  
29 and remediation schemes. In particular, models need to incorporate vertical variations in  
30 hydraulic conductivity in order to accurately simulate and predict solute transport.

31 Whilst horizontal variations can be characterized through hydraulic (pumping) tests, data  
32 on vertical variations are often lacking or have to be inferred from lithological or  
33 geophysical logs, because direct approaches such as packer testing are time consuming  
34 and costly. Fortunately, over the past two decades borehole flow logging techniques such  
35 as impeller logging have been developed which can quantify vertical hydraulic  
36 conductivity variations quickly and with relatively lightweight equipment. In outline, a  
37 well is pumped from the top and the variation in flow velocity with depth is measured  
38 using an impeller sonde, giving an indication of the location and magnitude of inflows to  
39 the well. However, approaches to the interpretation of flow logs have remained rather  
40 subjective, relying on visual inspection to distinguish meaningful data from noise. This  
41 paper presents a rigorous yet practical method for interpreting impeller logs to obtain a  
42 hydraulic conductivity profile with depth.

43

## 44 *Overview of impeller flow logging*

45 The impeller flow logging method has been described by several authors, including Jones  
46 and Skibitzke (1956), Schimschal (1981), Molz et al (1989, 1994), Paillet (1998, 2000)

and Gossell et al (1999). A pump is positioned in the casing at the top of the borehole, and run for sufficient time to attain quasi-steady state flow.

In order to flow log a well, the impeller sonde is lowered down the uncased section while the well is pumped at a constant rate from the top. The impeller flowmeter has a spinner whose rotation rate is related to vertical fluid flow speed. The flow rate recorded decreases progressively as each producing zone is passed. If there is no ambient flow in the well, the first derivative of flow speed is in theory related to hydraulic conductivity i.e. the impeller log identifies the relative contribution of each layer to overall transmissivity. Absolute hydraulic conductivity can be determined where aquifer transmissivity is known (Molz et al 1989, Keys 1990). The basic principle of the technique is illustrated in figure 1.

Although the impeller flowmeter has a low resolution compared to the heat-pulse flowmeter (Paillet 1998), flow can be logged while the tool is trolling (moving up or down the borehole). This is quicker than taking a series of static measurements, so in a given timeframe, more measurements can be made, giving a higher spatial resolution. In addition, the flow speed that the impeller “sees” during a static measurement is less than during a trolled measurement, and hence may be below the stall speed of the impeller. Upper limits of flow measurement are typically greater than for other flowmeters, allowing measurements over a greater range of conditions. This paper could also be applied to electromagnetic flowmeters which can also be trolled (Molz et al 1994).

The impeller flowmeter is the most common flow logging approach used in commercial borehole geophysical surveys. The trolling technique is typically used. Data quality may be limited because equipment is not of the highest specification, or for practical reasons associated with the borehole (e.g. partially collapsed zones). The technique presented in this paper will be of use to anyone who deals with impeller flow logs, but will be of particular use to practitioners trying to extract useful information from sub-prime datasets.

### *Interpreting flow logs*

Data quality may be limited in flow logs because of:

1. Turbulent flow causing high frequency fluctuations in velocity. Turbulence results from rugosities in the borehole wall (Morin et al 1988) and a high flow rate. Turbulence is greatest near the wall, so sonde centralisation is crucial (Keys 1990). Turbulence is also caused where fluid enters the borehole horizontally (Leach 1974).
2. Variations in borehole diameter (Tsang et al 1990, Pedler 1992, Paillet 2004). For example, although it might be expected that flow speed should fall downstream of a diameter increase, in reality the flow may “jet” into the enlarged hole section so that the expected drop in flow velocity is not seen (Bearden 1970, Hill 1990).
3. The construction of all but the most modern impellers is such that they can only measure whole numbers of rotations. This is a major limitation on their precision, especially at lower flow rates.

91 4. A non-linear relationship between spinner response and flow rate, particularly at low  
92 flow rates. This was found in calibrations performed by Hill (1990). Typically, a linear  
93 response is assumed, which is a source of error.

94 5. Changes in trolling speed. Even if the winch feeds out the cable at a steady rate, the  
95 centralizers cause the probe to slide down the borehole at an unsteady rate, especially in  
96 wells with ledges and washout zones.

97  
98 Taking the gradient of a raw flow log using finite differences between adjacent data  
99 points tends to amplify problems with data quality. To avoid these problems, workers  
100 typically interpret a flow log by manually identifying inflows where the flow log gradient  
101 is greatest and hence the hydraulic conductivity is highest (eg Schimschal 1981, Morin et  
102 al 1988, Paillet 1998, Gossell et al 1999, Sukop 2000). Such an approach incorporates a  
103 high degree of subjectivity and makes error estimation very difficult. Although it is  
104 sometimes possible to define the upper and lower limits of any inflowing horizons by the  
105 position of well screens or lithological contrasts ascertained from geophysical logs  
106 (Hanson and Nishikawa 1996), variation in permeability within the defined intervals is  
107 lost. The work of Fienan et al (2004) is a significant attempt at solving this problem,  
108 however, whilst this method is excellent for extracting small scale variations in zones of  
109 approximately constant hydraulic conductivity, it requires the zone boundaries and  
110 average permeability as an input, which may not be known parameters.

111  
112 Here, a discrete layer modeling approach has been developed which is aimed at  
113 overcoming the problem of noise in flow logs. It provides an objective and quantitative

analysis of flow variation in terms of hydraulic conductivity, together with a calculation of the uncertainty of the results presented. This method is illustrated through examples from the Chalk aquifer of East Yorkshire.

## **Method**

### *Theoretical basis*

Javandel and Witherspoon (1969) show that under quasi-steady state conditions the flow from an aquifer layer is proportional to its hydraulic conductivity. At the pumping rate normally required for impeller flow logging, the pressure gradient is the same for all producing zones, and the horizontal pressure gradient (i.e. drawdown) is much larger than any existing vertical gradient.

Using notation defined in figure 2, for an infinitesimally thin ‘layer’ of the aquifer,

$$dQ = -\alpha k dz \quad (1)$$

where  $dQ$  is the net flow from a layer,  $\alpha$  is a constant of proportionality,  $k$  is the hydraulic conductivity, and  $dz$  is the layer thickness. Using a method based on that of Molz et al (1989) and Fienan et al (2004) the constant of proportionality can be determined by integrating over the length of the well below the pump inlet:

$$\int_{Q_{\max}}^0 dQ = -\alpha \int_{z_{\min}}^{z_{\max}} k dz = -\alpha T \quad (2)$$

$$\therefore \alpha = \frac{Q_{\max}}{T} \quad (3)$$

where  $T$  is transmissivity, obtained from pump tests and  $Q_{\max}$  is the pump rate,  $z_{\max}$  is the depth of the borehole and  $z_{\min}$  is the depth at the top of the flow log (just below the pump

inlet). Assuming that the flow velocity in the borehole,  $v$ , is proportional to  $Q$  at all depths (i.e. provided borehole diameter variations are corrected for where necessary), and substituting for  $\alpha$  (equation 3) in equation (1), we obtain:

$$\frac{dv}{dz} = -\frac{v_{\max}}{T} k \quad (4)$$

so hydraulic conductivity is proportional to the first derivative of flow velocity with borehole depth.

#### *Practical considerations*

Impeller flowmeters are normally logged when the tool is trolling downwards, as this maximises impeller speed and the body of the tool does not shield the impeller (Hill 1990). The impeller is usually fitted with a mechanical centraliser to ensure it is kept central within the borehole. The values recorded during the logging process are the rotation rate of the impeller (in rotations per minute, RPM). In order to convert these values into the speed of fluid in the well, the sonde is calibrated using, for example, the technique of Leach et al (1974) and Syms et al (1982). In this technique the tool is trolled down a borehole containing static water (for example in the cased part of the borehole) and the average RPM,  $\omega$  is recorded at a variety of line speeds. The RPM,  $\omega$  is then plotted against line speed ( $l$ ) to give a response slope ( $m$ ) and a threshold value ( $v_{\omega=0}$ , the intercept with the line speed axis). The threshold value is the lowest velocity required to start the spinner rotating (i.e. to overcome friction in the bearings). Fluid velocity ( $v$ ) can then be found using the following equation:

158 
$$v = \frac{\omega}{m} + v_{\omega=0} - l \quad (5).$$

159

160 **Study Area**

161 The Cretaceous Chalk aquifer provides about 20% of the United Kingdom's drinking  
162 water supply (UK Groundwater Forum, 1998). Although the Chalk has a high matrix  
163 porosity of approximately 35% (e.g. Hartmann et al 2007), the pore throats are small (0.1  
164 to 1µm), resulting in a low effective matrix hydraulic conductivity of  $\sim 10^{-4}$  m/day (Price  
165 et al 1993). Bulk permeability is provided through flow in fractures, along bedding  
166 planes, joints normal or steeply inclined to the bedding planes, and faults (Patsoules  
167 1990). Many fractures are enlarged by dissolution (Price 1987). In places the Chalk is  
168 confined by a layer of Quaternary deposits consisting of low permeability glacial tills and  
169 alluvial organic clays.

170

171 Hydraulic conductivity varies with depth in the Chalk for three principle reasons:

172 1. Fracture frequency varies with depth (this is observed in cliff sections and cores).  
173 2. Marl and flint layers present within the Chalk act as aquitards, and layers of high  
174 hydraulic conductivity develop above them owing to solutional processes (Allshorn et al,  
175 2007).  
176 3. Periglacial weathering has caused further fracturing of the chalk at depths up to several  
177 tens of metres from the paleo-land surface (Higginbottom and Fookes, 1970).

178

179 The three example datasets presented here are from wells located at Benningholme  
180 (latitude: 53.8343N, longitude: 0.294410W) and North End Stream (latitude: 54.0115N,



longitude: 0.4419W) and Carnaby (latitude 54.0662N, longitude 0.2427W), in East Yorkshire in the UK. Logging at Benningholme and North End Stream was carried out by the University of Leeds, whereas the Carnaby logging was carried out by the British Geological Survey. At Benningholme, the confining layer is 16m thick, and at Carnaby it is 19m thick (Bloomfield and Shand, 1998). At North End Stream the aquifer is unconfined. The piezometric level is very close to the ground surface at all sites. The Benningholme and Carnaby wells are cased through the glacial till and all the wells are cased through the upper, most weathered part of the Chalk. Below this horizon they are open to the aquifer with no well screen or gravel pack; this design is similar to the water abstraction boreholes in the region. At Benningholme, transmissivity from a five hour pump test using the Cooper Jacob method was estimated at  $52\text{m}^2/\text{day}$ . At North End Stream, transmissivity from a two hour pump test using the Theim method was estimated at  $1070\text{m}^2/\text{day}$ . No direct measurements of transmissivity were made at Carnaby, so for the purposes of this analysis, a transmissivity value of  $131\text{m}^2/\text{day}$  was taken from a nearby borehole. Unpumped runs of the impeller were performed at North End Stream and Carnaby, and there was found to be no ambient flow, whilst the location of Benningholme on the confined part of the Chalk suggests there is low ambient flow.

## **Results**

Impeller flow logs and caliper logs were recorded at Benningholme in summer 2006, North End Stream in spring 2007 and at Carnaby in winter 1996. These flow logs are shown in figure 3, and the caliper logs in figure 4. As the average well bore diameters were different at each borehole, (0.21m at Benningholme and 0.16m at North End

Stream) a unique calibration was used for each. The response slope and threshold values used were 0.101 rotations per minute and 2.65 m/min at Benningholme and 0.128 rotations per minute and 0.72 m/min at North End Stream. At Carnaby no calibration had been performed, however, the maximum velocity could be calculated from the pumping rate, and it was assumed that the flow rate was zero at the bottom of the well. The flow log was then scaled between these two extremes, according to the technique of Leach et al (1974) and Hill (1990).

The Benningholme data show a continuous, steady decrease in flow velocity from just below the casing (20m bgl (below ground level)) to 30m bgl, with a more gradual decrease below with flow speed reaching zero at approximately 62m bgl. The North End Stream data show an extremely rapid decrease just below the casing at just over 12m bgl. The rate of decrease slows with depth, and flow speed reaches approximately zero just above the bottom of the hole at 23m. At Carnaby again there is a rapid decrease just below the casing (26m bgl), followed by a steady decrease to 95m where another rapid decrease occurs. At Benningholme in particular, because there is a low flow rate, the lack of precision in the data is apparent. At Benningholme the pump inlet was at 10m bgl and the pump rate was  $0.09\text{m}^3/\text{min}$  giving a flow speed of 2.6 m/min in the casing. At North End Stream the pump inlet was at 8m bgl and the pump rate was  $0.3\text{m}^3/\text{min}$  giving a flow speed of 16 m/min in the casing. The flow speed values just below the pump correspond closely to those predicted from the pumping rate in each case, which demonstrates that the impeller flowmeter calibration is accurate and the impeller is performing well. The caliper logs show that borehole diameter variations are relatively

small with only 10% variation from the average at Benningholme, 0.5% at North End Stream and 25% at Carnaby, so no corrections for borehole diameter were used when interpreting flow velocities.

### *Filtering*

It is clear from the data scatter in figure 3 that a simple finite difference method will produce hydraulic conductivity profile which is so noisy it cannot be usefully interpreted. Therefore, in order to apply the conventional interpretation for comparative purposes, an attempt to reduce the noise in the first derivative of flow speed was made using a smoothing algorithm. The Savitzky-Golay algorithm (Savitzky and Golay 1964, Press et al 1992) was chosen as it removes high frequency noise whilst preserving the magnitude and width of peaks in the data better than other techniques such as a running average. The algorithm performs a least squares fit to a polynomial on a window of consecutive data points and takes the calculated central point of the fitted polynomial curve as the new smoothed data point. The results are shown in figure 5 where the window used was 1 m with second order polynomials. The Benningholme plot shows a major peak at around 25m bgl, with other minor peaks below but all are interspersed with significant non-physical negative excursions. The North End Stream plot shows a smooth major peak at 11m bgl, but this is smeared to such an extent that it shows an unrealistic proportion of the hydraulic conductivity within the cased part of the hole. The Carnaby plot shows a significant peak at 26m bgl (again with an unrealistic proportion of the hydraulic conductivity within the cased part of the hole), with other minor peaks at 53m, 73m and 80m. In all cases, the algorithm smoothed the sections

where the gradient of the flow log changes, so there are gradual changes in hydraulic conductivity where the raw data suggests more discrete changes.

### **Discrete layer model**

An alternative approach to interpreting the flow rate curve is to define a physical model for hydraulic conductivity variation with depth, predicting the flow response of the model and matching it to the data using regression. Firstly a series of initial models consisting of various numbers of discrete layers are defined. The hydraulic conductivity of each layer and its thickness are optimised to provide the best match to the data using a least squares regression approach. If the layer boundaries are known from other geophysical logs or knowledge of the borehole construction they can be defined at this stage. Then, if the most likely model is not apparent from previous knowledge, a maximum likelihood approach is used to determine the most appropriate number of discrete layers present within the defined section. This approach has been applied to the flow logs from the Benningholme, North End Stream and Carnaby wells.

#### *The discrete layer model*

In the discrete layer modelling approach it is assumed that:

1. The analysis section consists of discrete layers each with a constant hydraulic conductivity.
2. The well sections to be analysed are bounded by layers with a hydraulic conductivity of zero. The upper layer would typically be inside the casing where there are no inflows, and the lower layer would be below the lowest inflow. The

recommended approach adopted here is to define the top of the uppermost  
conductive layer as the base of the casing and allow the modelling algorithm to  
identify the lowest inflow.

3. The horizontal pressure gradient generated by pumping is much larger than any  
existing vertical gradient. The analysis for cases where this does not hold is  
described at the end of this section.

Note that assumptions 1 and 2 could be altered to reflect any conceptual model of vertical  
hydraulic conductivity variation. For each borehole, the hydraulic conductivity,  $k(z)$ , of  
the aquifer at depth  $z$  is defined as:

282

$$k = \begin{cases} 0 & z \leq z_0, \\ \dots & \dots \\ k_i & z_{i-1} < z \leq z_i, i = 1..n \\ \dots & \dots \\ 0 & z > z_n \end{cases} \quad (6)$$

where  $k_i$  are constant hydraulic conductivities in  $n$  layers of the aquifer and  $z_i$  are the  
depths of the layer interfaces, with  $z_0$  being the depth of the base of the solid casing.

286

287 *Modelling the flow rate*

288 The flow rate in the borehole,  $v(z)$  during pumping can be inferred using equation (4):

$$\frac{v(z)}{v_{\max}} = 1 - \frac{\int_0^z k dz}{T} = \begin{cases} 1 & z \leq z_0, \\ \dots & \dots \\ 1 - \frac{1}{T} \left( k_i (z - z_{i-1}) + \sum_{j=1}^{i-1} k_j (z_j - z_{j-1}) \right) & z_{i-1} < z \leq z_i, \\ \dots & \dots \\ 1 - \frac{1}{T} \sum_{j=1}^n k_j (z_j - z_{j-1}) & z > z_n \end{cases} \quad (7)$$

290 where  $v_{\max}$  is the maximum flow speed at the top of the borehole (within the casing) and  
 291  $T$  is the overall transmissivity (e.g. found from pumping tests).

292

293 This discrete layer model has abrupt transitions of the first derivative of flow velocity at  
 294 the interface between conductive layers which makes fitting to data problematic for  
 295 regression algorithms. This problem is eased by smoothing the abrupt transitions, using a  
 296 method based on that of Bacon and Watts (1971). Here the modelled flow velocity is  
 297 multiplied by a notch function  $c(z)$  depths  $z_a$  and  $z_b$ :

$$298 \quad c(z, z_a, z_b, g) = \frac{1}{4} \left( 1 + \tanh\left(\frac{z - z_a}{g}\right) \right) \left( 1 - \tanh\left(\frac{z - z_b}{g}\right) \right) \quad (8)$$

299 The parameter  $g$  (units metres) is the distance from the transition where 12% of the effect  
 300 of the notch filter remains, that is,  $g$  defines the sharpness of the edges of the notch  
 301 function, as shown in figure 6. The model is then redefined as:

$$302 \quad \frac{v(z)}{v_{\max}} = \left[ 1 - \frac{1}{T} \left( k_i(z - z_{i-1}) + \sum_{j=1}^{i-1} k_j(z_j - z_{j-1}) \right) \right] \cdot c(z, z_{i-1}, z_i, g) + \dots$$

$$\left[ 1 - \frac{1}{T} \sum_{j=1}^n k_j(z_j - z_{j-1}) \right] \cdot c(z, z_n, \infty, g)$$

303 Simplifying and taking out the 1 that occurs in all segments of the function:

$$304 \quad \frac{v(z)}{v_{\max}} = 1 - \sum_{i=1}^{i=n} \left[ \frac{1}{T} \left( k_i(z - z_{i-1}) + \sum_{j=1}^{i-1} k_j(z_j - z_{j-1}) \right) c(z, z_{i-1}, z_i, g) \right] - \frac{1}{T} \left( \sum_{j=1}^n k_j(z_j - z_{j-1}) \right) c(z, z_n, \infty, g)$$

305 (10)

As  $g$  tends to zero, equation 10 tends to equation 7. It was found that  $g=0.1$  m was a good compromise providing minimal distortion to the function, whilst giving few problems for the regression algorithm.

#### *Fitting the data to the model*

The model function is non-linear, and the points of discontinuity (representing layer interfaces) are parameters to be determined (if they are unknown from other geophysical logs or the borehole construction), so this model is non-linear and requires solution by iterative means (e.g. Bates and Watts, 1988). The regression is performed using the ‘nls’ function in the statistical analysis software ‘R’ (R. Development Core Team, 2007). Code for performing the regression in R is included in the online supplementary material. For these cases, it was possible to fix the  $z_0$  value at the bottom of the well casing. R also allows confidence intervals for each model parameter to be computed.

Several regression analyses were carried out on each data set using different values of  $n$ , the number of conductive layers. In theory, the number of layers,  $n$ , could increase up to a value one less than the number of data points. However, eventually the model will be fitting to noise and not to the real signal, for example when the model gives a physically impossible negative value of hydraulic conductivity.

It was found possible to generate five realistic models at Benningholme and North End Stream, and two at Carnaby. The values for  $v_{max}$ ,  $k_i$  and  $z_i$  for each model fitted to the data are given in tables 1 to 3. Selected model curves are shown graphically in figure 7,

329 along with the original data. At Benningholme (figure 7a) the  $n=3$  model shows two  
330 narrow layers at the top, the uppermost with a shallow gradient ( $dv/dz$ ) and the other with  
331 a steep gradient. The lowermost layer of the model has a very shallow gradient; this layer  
332 is thickest and actually represents most of the borehole. The  $n=4$  model is similar except  
333 that a third narrow layer is added below the previous two narrow layers; this layer has a  
334 gradient shallower than the second layer but steeper than the lowermost layer. The  $n=5$   
335 model is visually indistinct from the  $n=4$  model at the scale plotted in figure 7a, but  
336 inspection of table 1 reveals a very narrow layer with a steeper gradient is added to the  
337 base of the previous third layer.

338  
339 At North End Stream (figure 7b), the layers in the  $n=3$  model increase in thickness and  
340 decrease in gradient with depth. The  $n=4$  model is similar to the  $n=3$  model, but a layer  
341 with steep gradient is added at the bottom of the lowermost layer. The  $n=5$  model is  
342 indistinguishable from the  $n=4$  model in figure 8b, but inspection of the data (table 2)  
343 shows the former third layer has been divided in half, with the upper of these two layers  
344 showing a slightly steeper gradient than the lower.

345  
346 At Carnaby (figure 7c) the one layer in the  $n=1$  model has a single layer with a very steep  
347 gradient that occurs just below the casing. Below this, the model is clearly a poor fit to  
348 the data. The  $n=3$  model takes the layer from the  $n=1$  model as its top layer, then adds a  
349 wide layer with a shallow gradient, and another steep gradient layer at the bottom. There  
350 are fewer models for this dataset than for the others. The regression analysis could not  
351 define a model for the  $n=2$  case, as assuming the top layer was similar to that for the  $n=1$



model, and the lower layer was the best fit to the data below, the lower bound of this layer would be below the bottom of the borehole. In the  $n=4$  and above models, unrealistic negative values for hydraulic conductivity were given for at least one layer.

At this stage in the analysis, there are still several different possible models, all with different numbers of layers. In this case, no logs other geophysical logs were available to help determine the number of layers in the most likely model, except for the caliper logs (figure 4). Benningholme and North End Stream do not show any fracture related widenings at the same location as the flowing fractures indicated by the models. The widest part of the Carnaby borehole is immediately below the casing where both models show a significant inflow. There is also a widening on the caliper log at 95m where the  $n=3$  model indicated a significant inflow. However, there are many widening on all three caliper logs which are not associated with flowing fractures. This suggests that for the Chalk, it is not sensible to use the caliper logs to inform the location of the layer boundaries. If suitable alternative logs were available (and there is confidence in their ability to accurately indicate layer boundaries) they could have been used to define the position of the layer boundaries, reducing the number of candidate models for the borehole hydraulic conductivity profile.

In this case there are multiple candidate models and a statistical criterion is required to allow valid comparisons to be made between models. If the root mean squared error ( $S_n$ ) for each model is considered, it will simply decrease as the  $n$  value increases, as the situation where the model merely joins each data point with a line is approached ( $S_n=0$ ).

What is required here is to identify the model that gives the optimum combination of minimized error value and model simplicity.

Akaike's Information Criterion (AIC) is a standard method for model comparison and selection in time series analysis and is gaining widespread acceptance in the biological sciences (Motulsky and Christopoulos 2003). AIC is an estimate of the relative information loss caused by approximating reality to a model (information being measured by the Kullback-Leibler divergence). The derivation of the model is rooted in information theory and Fisher's maximum likelihood (Burnham and Anderson, 2002). Once AIC is computed, it is possible to derive an estimate of the relative likelihood of each model.

The generic AIC makes few assumptions and the method can be used to evaluate models under any data conditions. However, using the standard assumptions of regression (residuals distributed normally, independently and with zero mean), AICc can be computed as:

$$AICc = N \log(S_n^2) + \left( \frac{8n^2 + 20n + 12}{1 - n} \right) + 4n + 4 \quad (10)$$

where N is the number of flow measurements. This is a corrected version of AIC for small samples (Akaike 1974). As AICc is only a relative estimate of information loss, the absolute value has no meaning, and instead the differences in AICc values between different models are considered, for example with reference to the model with the lowest value ( $AICc_{min}$ ):

$$\Delta AICc = AICc_i - AICc_{min} \quad (11)$$

The likelihood of each model is proportional to  $e^{(-0.5\Delta AIC)}$ . It is convenient to normalise the likelihoods so that they sum to unity:

$$w_i = \frac{e^{-0.5\Delta AIC_i}}{\sum e^{-0.5\Delta AIC_c}} \quad (12)$$

where  $w_i$  are Akaike's weights and can be conceptualised as the probability that a given model is correct, assuming that all plausible models have been considered. It should be realized that Akaike's weights are merely an expression of the relative likelihood of alternative models given the data. In practice, there may be other factors that suggest that one model may be better than another. Where this information is certain, it should be used to constrain the regression process. Where the information available is open to alternative interpretations and especially where AIC does not provide an unambiguous answer, model selection will be more difficult. As ever, the worker will need to weigh the different forms of evidence including Akaike's weights, geophysical logs and what is known of the local hydrogeology to draw appropriate conclusions.

Akaike's weights are reported in the last column of tables 1 to 3. These results suggest that for Benningholme and North End Stream the  $n=4$  model is the most likely, whilst at Carnaby the  $n=3$  model is most likely. This result seems credible based on visual inspection of the data, and there was no other information available to help select the model.

The values used for the best fit hydraulic conductivity profile are indicated in bold on tables 1 and 2, and the plotted in figure 8. The results from the modeling at

Benningholme (figure 8a) show the second layer from the top is very narrow yet has by far the highest hydraulic conductivity – it is probable that this layer represents flow from a single fracture. The top and third layers have very similar thicknesses and hydraulic conductivity values, whilst the lowest and thickest layer has a lower value of hydraulic conductivity. At North End Stream (figure 8b) the three uppermost hydraulic conductivity layers decrease in magnitude whilst increasing in thickness as depth increases. The lowermost layer is relatively thin and has a similar hydraulic conductivity to the second layer (although there is more uncertainty in both its thickness and its magnitude). At Carnaby the top and bottom layers are both very thin, and have comparable hydraulic conductivities, both significantly higher than the middle layer. The bottom layer has a very wide confidence limit – this is because there are very few data points that occur within this layer.

#### *Cases with ambient flow*

In applying the method to the flow logs described here it was assumed there was no ambient flow in the borehole. Although this is a reasonable assumption for the boreholes presented here, significant ambient flow would affect the flow logs and hence the application of the analysis method. There is a technique for combining the flow logs measured at two different pumping rates to obtain a hydraulic conductivity profile for a borehole with significant ambient flow, described, for example, by Paillet (2000). If hydraulic conductivities are converted into actual flow rates into or out of the borehole from an aquifer layer (negative hydraulic conductivities should of course be permitted as these represent flows leaving the borehole), and the borehole diameter, radius of

influence of the producing zone and water level difference between the two different pumping regimes are known, a hydraulic conductivity profile can be calculated (for example from Paillet 2000, equation 2), even if there are significant ambient flows.

## **Discussion**

A new method for analysing flow logs has been presented. Conceptual models are defined, consisting of layers of constant hydraulic conductivity. Each model has a different number of hydraulically conductive layers. The models are fitted to the flow log data, so that the depths of the top and bottom of each conductive layer and its hydraulic conductivity are defined. If necessary, the models are then tested against one another using a simplicity-versus-fit test to establish which is the most likely representation of the data, to avoid over-fitting. For the most likely model, confidence intervals of the depths of layer interfaces and the magnitudes of the hydraulic conductivity are then calculated. In this way, it is possible to take a raw flow log calculated from an impeller sonde, and convert it into a plot of hydraulic conductivity against depth. Hence a user of the data could, for any given depth, read off the value of hydraulic conductivity and have an idea of the certainty of that value.

In the examples presented here, the automated techniques used produced satisfactory results. According to the confidence limits, the depths of the layer boundaries are defined to within  $\pm 0.6\text{m}$  (and usually only a few centimetres). As figure 9 shows, the confidence limits for the hydraulic conductivity values are also typically small except

when thin layers are present in which case uncertainty in layer boundary position translates into a larger uncertainty in hydraulic conductivity.

The modelling approach used has indicated that the flow into boreholes within the Chalk aquifer consists of both major contributions from individual flowing fractures and layers of more distributed hydraulic conductivity which is approximately constant over depth intervals of several metres. The hydraulic conductivity in these layers is still likely to result from fracture flow, but with individual fracture contributions below the resolution of the flowmeter. This is consistent with the structure of the Chalk aquifer (see section 3). Where fracture frequency is high, the distributed hydraulic conductivity will also be high. Periglacial weathering has increased fracture frequency in the Chalk nearest the land surface which is consistent with the cases presented here where the hydraulic conductivity is concentrated near the top of the aquifer. In addition, marl and flint layers form aquitards, so individual fractures above them will become solutionally enhanced and will contribute significantly to borehole flow.

This modelling method is a significant improvement on filtering, as it can pick out individual contributing fractures. The filtering process shows peaks in the correct places, but cannot distinguish between the distributed zones and individual fractures. It may be possible to pick out these fractures by manual study of the data, but the modelling process can do this rigorously.

The pumping rates used in the tests presented here were high ( $0.09\text{m}^3/\text{min}$  at Benningholme,  $0.3\text{m}^3/\text{min}$  at North End Stream and  $0.4\text{m}^3/\text{min}$  at Carnaby). These rates were used because it was found that with lower pump rates, the impeller often stalled. The impeller used in this study was not fitted with a diverter or packer because this disturbs the flow profile being measured (Ruud et al 1999, Zlotnik and Zurbuchen 2003). Any variations in pump rate would contribute to the noise in the data. The logging was carried out after quasi-steady state had been reached, so there was no systematic variation in flow rate related to increasing drawdown. Measurements taken during logging suggested that pumping rate varied by only  $\pm 7\%$  from the mean, so it is not thought that variations in pump rate contribute significantly to data noise.

## Conclusions

An automated method for interpreting well impeller flow log data by fitting a model consisting of several layers with discrete hydraulic conductivity values has been developed. The discrete layer model fitting technique picks up subtleties in the data that might have gone unnoticed in manual interpretation, including the ability to delineate two or more adjacent layers of similar hydraulic conductivity. The model fitting approach also represents a rigorous method of defining the depth of the layer boundaries, hydraulic conductivities and the uncertainty in these key parameters. It can be used in situations where there is no knowledge of lithological changes with depth, or in rocks like the chalk where the lithological contrasts that define high hydraulic conductivity layers are not detected by geophysical logs.

The proposed method represents a significant improvement to the presently commonly used methods for interpreting flow logs, as it involves less human intervention than conventional manual approaches, but avoids the spurious negative values of hydraulic conductivity produced by numerical differentiation. The techniques used (non-linear regression, the notch function, Akaike's Information Criterion) can all be implemented in freely available software, making the method easily accessible to all potential users.

## **Appendix**

The supplementary online material contains code for the freely available software "R", which will enable the user to perform the regression described here and calculate confidence intervals.

## **Acknowledgements**

The authors would like to acknowledge financial support for this work from the Natural Environment Research Council Grant Number NER/S/A/2005/13328 at the University of Leeds, and from the Environment Agency NE Region (Mr. Rolf Farrell). The fieldwork would not have been possible without the co-operation of Claire Binnington and Chris Brotherton at Drifffield Town Council, Mr Harrison at Wilfholme Landing, and Jonathan Basing from the Environment Agency. The BGS kindly allowed access to the data collected at Carnaby. We would also like to thank the late Dr Steve Truss for field assistance and Kirk Handley for technical assistance. Fred Paillet, Michal Pitrak and an anonymous reviewer are also thanked for their comments on the manuscript.



533 **References**

534 Akaike, H. 1974. A new look at the statistical model identification, IEEE Transactions on  
535 Automatic Control 19, no 6: 716–723.

536 Allshorn, S.J.L., S.H. Bottrell, L.J. West and N.E. Odling. 2007. Rapid karstic bypass  
537 flow in the unsaturated zone of the Yorkshire chalk aquifer and implications for  
538 contaminant transport, in: Natural and Anthropogenic Hazards in Karst Areas:  
539 Recognition, Analysis and Mitigation, edited by M. Parise, J. Gunn, Geological Society  
540 Special Publication 279: 111-122, Geological Society of London, UK.

541 Bacon, D.W. and D.G. Watts.1971. Estimating the transition between two intersecting  
542 lines, Biometrika 58, no 3: 525–534.

543  
544 Bates D.M. and D.G. Watts. 1988. Nonlinear regression analysis and its applications,  
545 New York: John Wiley and Sons

546  
547 Bearden, W.G., D. Currens, R.D. Cocanower, and M. Dillingham 1970. Interpretation of  
548 injectivity profiles in irregular bore holes, Journal of Petroleum Technology 22, no 9:  
549 1089-1097.

550  
551 Bloomfield, J.P. and P. Shand. 1998. Summary of the Carnaby Moor Borehole  
552 investigation, British Geological Survey Technical Report WD/98/30.

553 Burnham. K.P. and Anderson, D. 2002. Model Selection and Multi-Model Inference, 2<sup>nd</sup>  
554 edition, New York: Springer

555 Gale, I.N. and H.K. Rutter. 2006. The Chalk aquifer of Yorkshire, Research Report  
556 RR/06/04, British Geological Survey, Nottingham, UK

557  
558 Gossell M.A., T. Nishikawa, R.T. Haqnsen, J.A. Izbicki, M.A. Tabidian and K. Bertine.  
559 1999. Application of flowmeter and depth-dependent water quality data for improved  
560 production well construction, Ground Water 37, no 5: 729–735.

561  
562 Hanson R.T., and T. Nishikawa. 1996. Combined use of flowmeter and time-drawdown  
563 data to estimate hydraulic conductivity in layered aquifer systems, Ground Water 34, no  
564 1: 84-94.

565  
566 Hartmann, S., N.E. Odling and L.J. West. 2007. A multi-directional tracer test in the  
567 fractured Chalk aquifer of E. Yorkshire, UK, Journal of Contaminant Hydrology, 94, no  
568 3-4: 315-331.

569  
570 Higginbottom, I.E., and P.G. Fookes. 1970. Engineering aspects of periglacial features in  
571 Britain, Quarterly Journal of Engineering Geology and Hydrogeology 3, no 2: 85-117.

572  
573 Hill, A.D. 1990. Production logging—Theoretical and interpretive elements, Society of  
574 Petroleum Engineers Monograph 14, Richardson, Texas: Society of Petroleum Engineers,

575  
576 Javandel, I., and P.A. Witherspoon. 1969. A method of analyzing transient fluid flow in  
577 multilayered aquifers, *Water Resources Research* 5, no 4: 856–869  
578  
579 Jones P.H., and E.B. Skibitzke. 1956. Subsurface geophysical methods in ground water  
580 hydrology, *Advances in Geophysics* 3: 241-297  
581  
582 Keys, W.S. 1990. Borehole geophysics applied to ground-water investigations. USGS  
583 *Techniques of Water-Resources Investigations Book 2*, Chapter E2.  
584  
585 Leach, BC, J.B. Jameson, J.J. Smolen and Y. Nicolas. 1974. The full bore flowmeter,  
586 SPE Paper number 5089  
587  
588 Molz, F.J., R.H. Morin, A.E. Hess, J.G. Melville, and O. Guven. 1989. The impeller  
589 meter for measuring aquifer permeability variations: Evaluation and comparison with  
590 other tests, *Water Resources Research* 25, no7: 1677-1683.  
591  
592 Molz, F.J., G.R. Boman, S.C. Young and W.R. Waldrop. 1994. Borehole flowmeters:  
593 field application and data analysis, *Journal of Hydrology* 163, no 3-4: 347-371  
  
594 Morin, R.H., A.E. Hess and F.L. Paillet. 1988. Determining the distribution of hydraulic  
595 conductivity in a fractured limestone by simultaneous injection and geophysical logging:  
596 *Ground Water* 26, no 5: 587-595.  
  
597 Motulsky, H. and A. Christopoulos. 2003. *Fitting Models to Biological Data Using*  
598 *Linear & Nonlinear Regression A Practical Guide to Curve Fitting*, San Diego,  
599 California: GraphPad Software Inc.  
600  
601 Paillet, F.L. 1998. Flow modelling and permeability using borehole flow logs in  
602 heterogeneous fractured formations, *Water Resources Research* 34, no 5: 997-1010  
603  
604 Paillet, F.L. 2000. A field technique for estimating aquifer parameters using flow log  
605 data, *Ground Water* 38, no 4: 510-521.  
606  
607 Paillet, F.L. 2004. Borehole flowmeter applications in irregular and large diameter  
608 boreholes, *Journal of Applied Geophysics* 55, no 1-2: 39-59.  
609  
610 Patsoules, M.G. 1990. Survey of macro and micro-fracturing in the Yorkshire Chalk, in:  
611 *Chalk: Proceedings of the International Chalk Symposium held at Brighton Polytechnic*  
612 *on 4-7 September 1989*, edited by J.B. Burland: 87-93,  
613  
614 Pedler, W.H., C.L. Head and L.L. Williams. 1992. Hydrophysical logging: a new  
615 wellbore technology for hydrogeologic and contaminant characterization of aquifers. In:  
616 *6th National. Outdoor Action Conference., Proceedings. National. Ground Water*  
617 *Association: 701–715.*  
618

- Press W.H., S.A. Teukolsky, W.T. Vetterling and B.P. Flannery. 1992. Numerical Recipes in FORTRAN, The Art of Scientific Computing, Second Edition, Cambridge, UK: Cambridge University Press,
- Price M.1987. Fluid flow in the Chalk of England, Fluid flow in sedimentary basins and aquifers, edited by J.C. Goff and B.P.J. Williams, Geological Society Special Publication No 34: 141-156, London, UK: Geological Society,
- Price M., R.A. Downing and W.M. Edmunds.1993. The chalk as an aquifer, The Hydrogeology of the Chalk of North-West Europe, edited by R.A. Downing, M. Price and G.P. Jones: 35-58, Oxford, UK: Oxford University Press,
- R Development Core Team. 2007. R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria.
- Ruud, N.C., Z.J. Kabala, Z.J., and F.J. Molz. 1999. Evaluation of flowmeter-head loss effects in the flowmeter test, Journal of Hydrology 224, no 1: 55–63.
- Savitzky, A. and M.J.E. Golay. 1964. Smoothing and Differentiation of Data by Simplified Least Squares Procedures; Analytical Chemistry 36, no 8: 1627-1639.
- Schimschal U. 1987. Flowmeter analysis at Raft River, Idaho, Ground Water 19, no 1: 93-97.
- Sukop, M.C. 2000. Estimation of vertical concentration profiles from existing wells. Ground Water 38, no 6: 836-841.
- Syms M.C., P.H. Syms, P.F. Bixley. 1982. Interpretation of flow measurement in geothermal wells without caliper data, Log Analyst 23, no 2: 34-45.
- Tsang, C., P. Hufschmied and F.V. Hale. 1990. Determination of fracture inflow parameters with a borehole fluid conductivity logging method, Water Resources Research 26, no 4: 561-576.
- UK Ground Water Forum. 1998. Groundwater: our hidden asset, Nottingham, UK: British Geological Survey,
- Zlotnik, V.A., B.R Zurbuchen. 2003. Estimation of hydraulic conductivity from borehole flowmeter tests considering head losses, Journal of Hydrology 281, no 1-2: 115–128

<b>n</b>		<b>v<sub>max</sub></b>	<b>k<sub>1</sub></b>	<b>k<sub>2</sub></b>	<b>k<sub>3</sub></b>	<b>k<sub>4</sub></b>	<b>k<sub>5</sub></b>	
		<b>m/min</b>	<b>m/day</b>	<b>m/day</b>	<b>m/day</b>	<b>m/day</b>	<b>m/day</b>	
1	Best fit	2.3	5.1					
2	Best fit	2.3	1.4	8.0				
3	Best fit	2.3	1.1	12	0.30			
4	<b>Best fit</b>	<b>2.3</b>	<b>2.5</b>	<b>340</b>	<b>3.0</b>	<b>0.25</b>		
	2.5% confidence limit	2.3	2.5	230	2.7	0.23		
	97.5% confidence limit	2.3	2.7	450	3.2	0.25		
5	Best fit	2.3	2.5	340	3.0	79	0.25	
		<b>z<sub>0</sub></b>	<b>z<sub>1</sub></b>	<b>z<sub>2</sub></b>	<b>z<sub>3</sub></b>	<b>z<sub>4</sub></b>	<b>z<sub>5</sub></b>	<b>Akaike's weights</b>
		<b>m</b>	<b>m</b>	<b>m</b>	<b>m</b>	<b>m</b>	<b>m</b>	<b>%</b>
1	Best fit	20.02	29.94					0
2	Best fit	20.02	22.69	28.40				0
3	Best fit	20.02	23.23	26.45	59.54			0
4	<b>Best fit</b>	<b>20.02</b>	<b>24.77</b>	<b>24.83</b>	<b>29.20</b>	<b>62.63</b>		<b>53.0</b>
	2.5% confidence limit		24.75	24.81	28.98	62.00		
	97.5% confidence limit		24.78	24.85	29.43	63.09		
5	Best fit	20.02	24.77	24.83	29.98	29.99	62.63	47.0

**Table 1:** Model results for Benningholme with Akaike's Information Criterion results (last column).  $v_{\max}$  is the maximum flow speed,  $k_i$  is the hydraulic conductivity of each layer,  $z_0$  is the depth of the top of the uppermost layer, and  $z_i$  is the depth of the bottom of each layer.

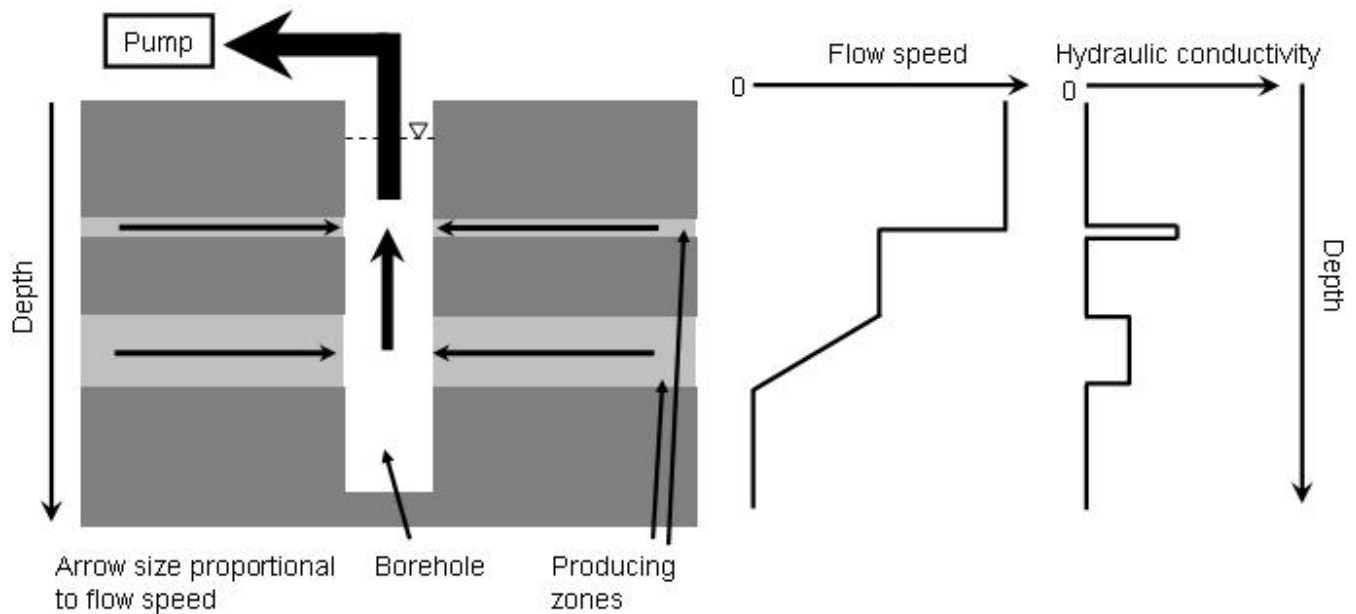
<b>n</b>		<b>v<sub>max</sub></b>	<b>k<sub>1</sub></b>	<b>k<sub>2</sub></b>	<b>k<sub>3</sub></b>	<b>k<sub>4</sub></b>	<b>k<sub>5</sub></b>	
		<b>m/min</b>	<b>m/day</b>	<b>m/day</b>	<b>m/day</b>	<b>m/day</b>	<b>m/day</b>	
1	Best fit	14	770					
2	Best fit	14	3400	46				
3	Best fit	14	4200	250	31			
4	<b>Best fit</b>	<b>14</b>	<b>4200</b>	<b>240</b>	<b>27</b>	<b>150</b>		
	2.5% confidence limit	14	3800	230	25	88		
	97.5% confidence limit	15	4400	240	27	420		
5	Best fit	14	4200	240	25	34	130	
		<b>z<sub>0</sub></b>	<b>z<sub>1</sub></b>	<b>z<sub>2</sub></b>	<b>z<sub>3</sub></b>	<b>z<sub>4</sub></b>	<b>z<sub>5</sub></b>	<b>Akaike's weights</b>
		<b>m</b>	<b>m</b>	<b>m</b>	<b>m</b>	<b>m</b>	<b>m</b>	<b>%</b>
1	Best fit	10.88	12.14					0
2	Best fit	10.88	11.08	19.28				0
3	Best fit	10.88	11.00	12.41	20.45			0
4	<b>Best fit</b>	<b>10.88</b>	<b>11.00</b>	<b>12.47</b>	<b>18.70</b>	<b>19.15</b>		<b>58.5</b>
	2.5% confidence limit		10.99	12.38	18.54	18.95		
	97.5% confidence limit		11.01	12.54	18.84	19.38		
5	Best fit	10.88	11.00	12.47	16.98	18.72	19.17	41.5

**Table 2:** Model results for North End Stream with Akaike's Information Criterion results (last column).

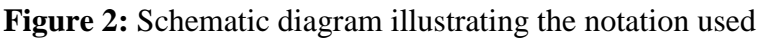
<b>n</b>		<b>v<sub>max</sub></b>	<b>k<sub>1</sub></b>	<b>k<sub>2</sub></b>	<b>k<sub>3</sub></b>	<b>z<sub>0</sub></b>	<b>z<sub>1</sub></b>	<b>z<sub>2</sub></b>	<b>z<sub>3</sub></b>	<b>Akaike's weights</b>
		<b>m/min</b>	<b>m/day</b>	<b>m/day</b>	<b>m/day</b>	<b>m</b>	<b>m</b>	<b>m</b>	<b>m</b>	<b>%</b>
1	Best fit	21	99			26.68	27.52			0
3	<b>Best fit</b>	<b>21</b>	<b>86</b>	<b>0.84</b>		<b>26.68</b>	<b>27.28</b>	<b>95.19</b>	<b>95.38</b>	<b>100</b>
	2.5% confidence limit	21	80	0.83	0.00	26.68	27.24	95.03	<sup>1</sup>	
	97.5% confidence limit	21	92	0.85	870	26.68	27.33	95.30	95.67	

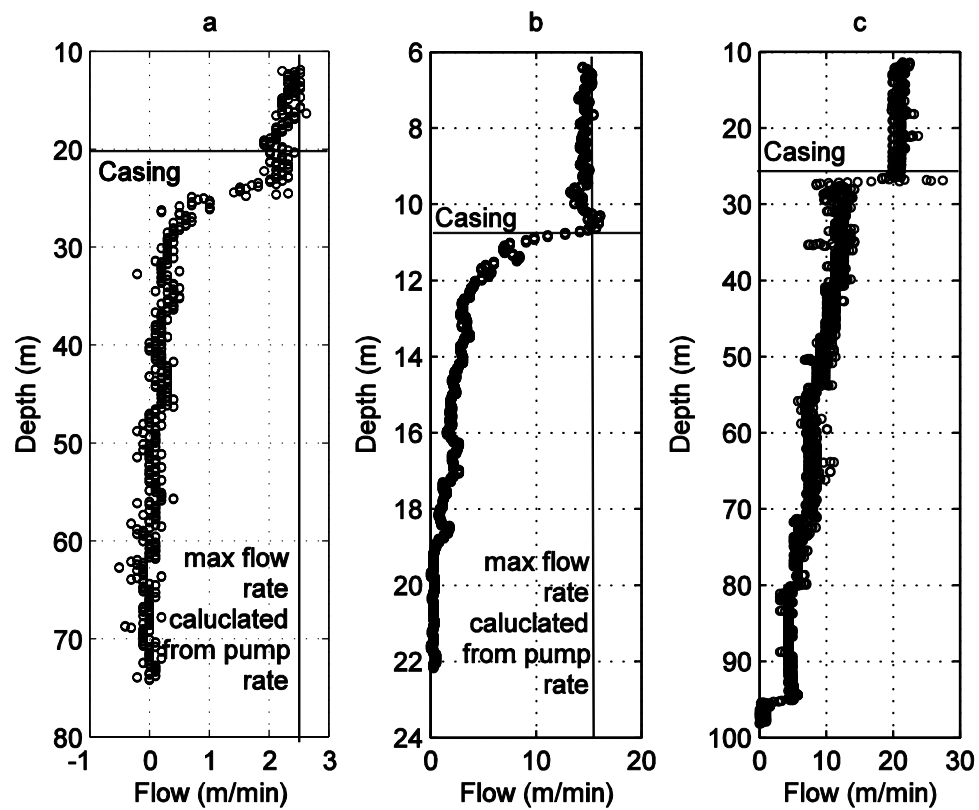
**Table 3:** Model results for Carnaby with Akaike's Information Criterion results (last column).

<sup>1</sup> Value not given because it overlaps with the 97.5% confidence interval for  $z_2$ .

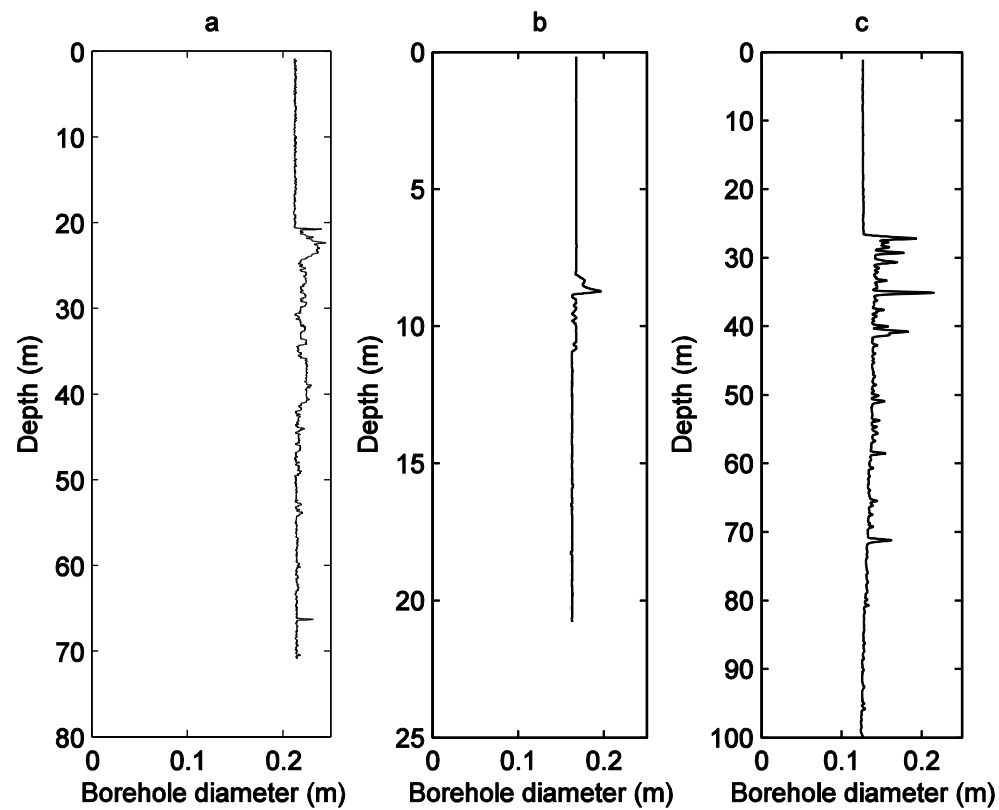


**Figure 1:** Schematic diagram illustrating the principles of using an impeller flow log to obtain a hydraulic conductivity profile



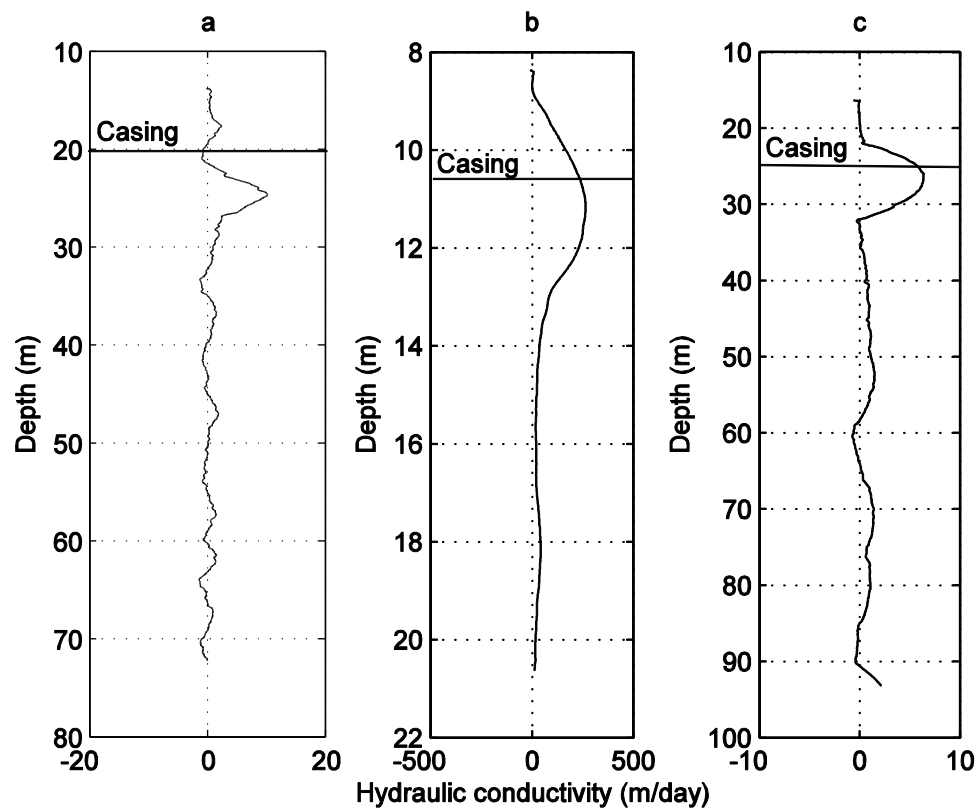


**Figure 3:** Flow profiles from a) Benningholme, b) North End Stream and c) Carnaby.

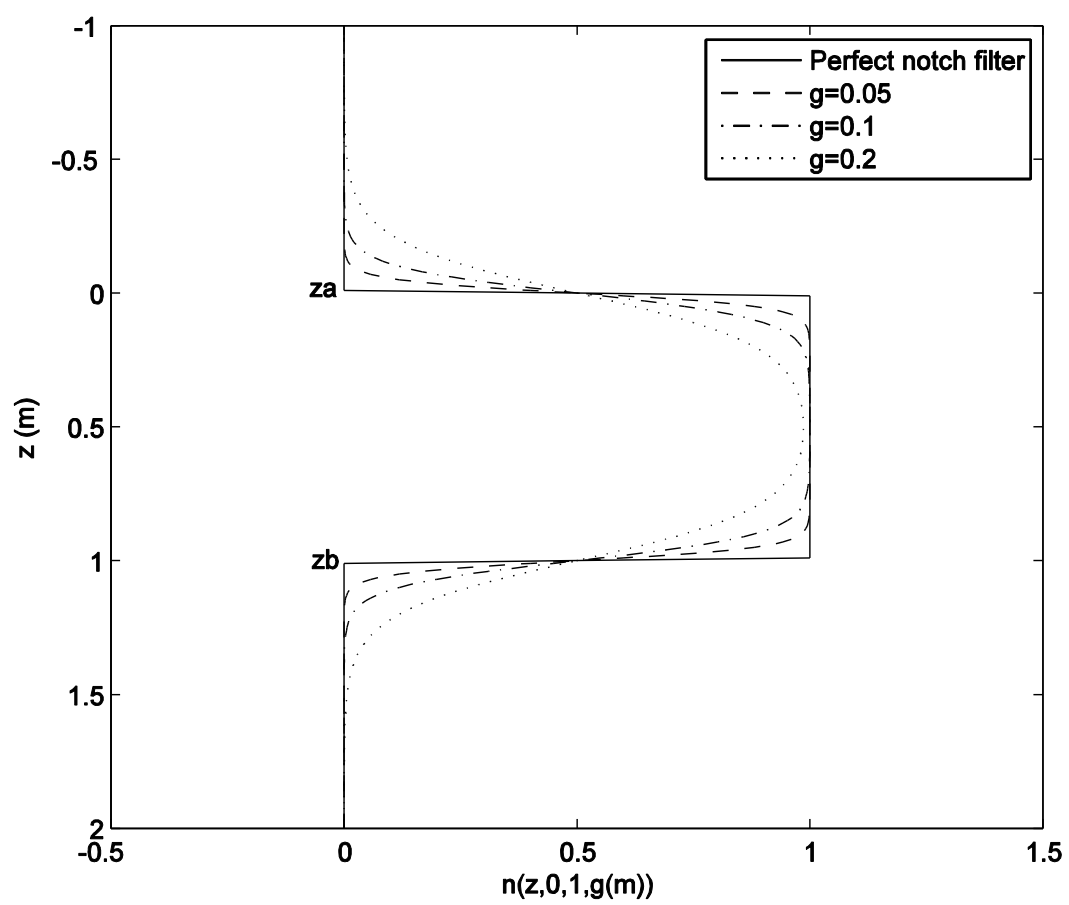


**Figure 4:** Caliper logs for a) Benningholme, b) North End Stream and c) Carnaby

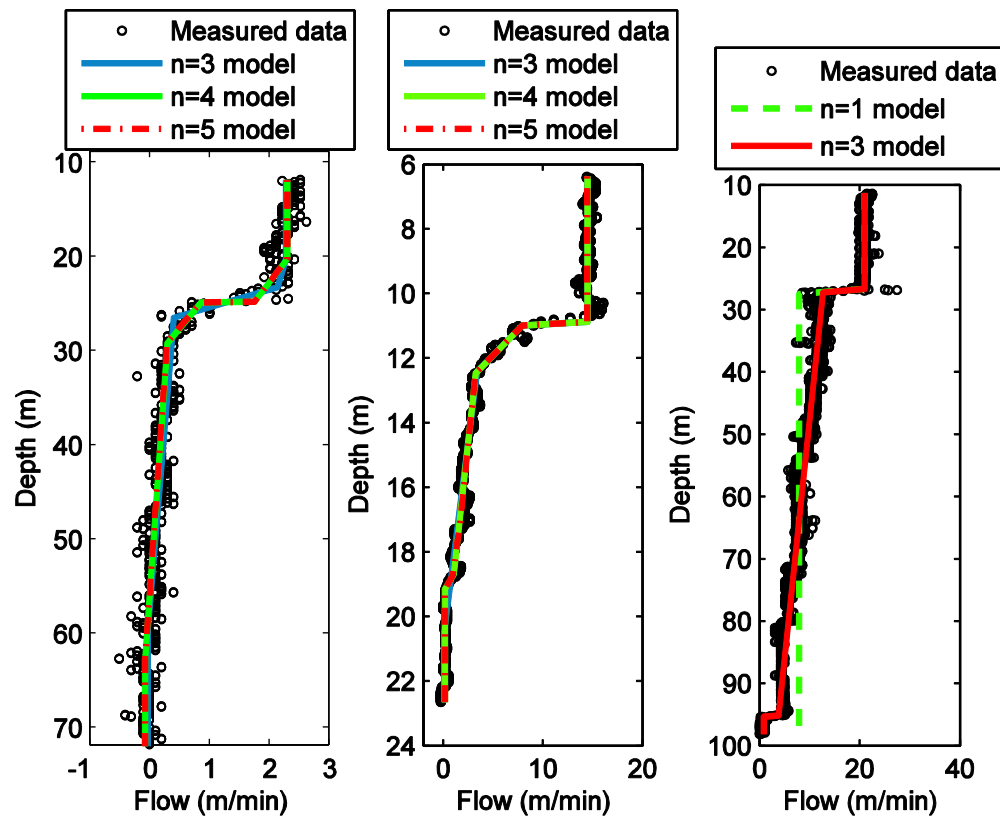




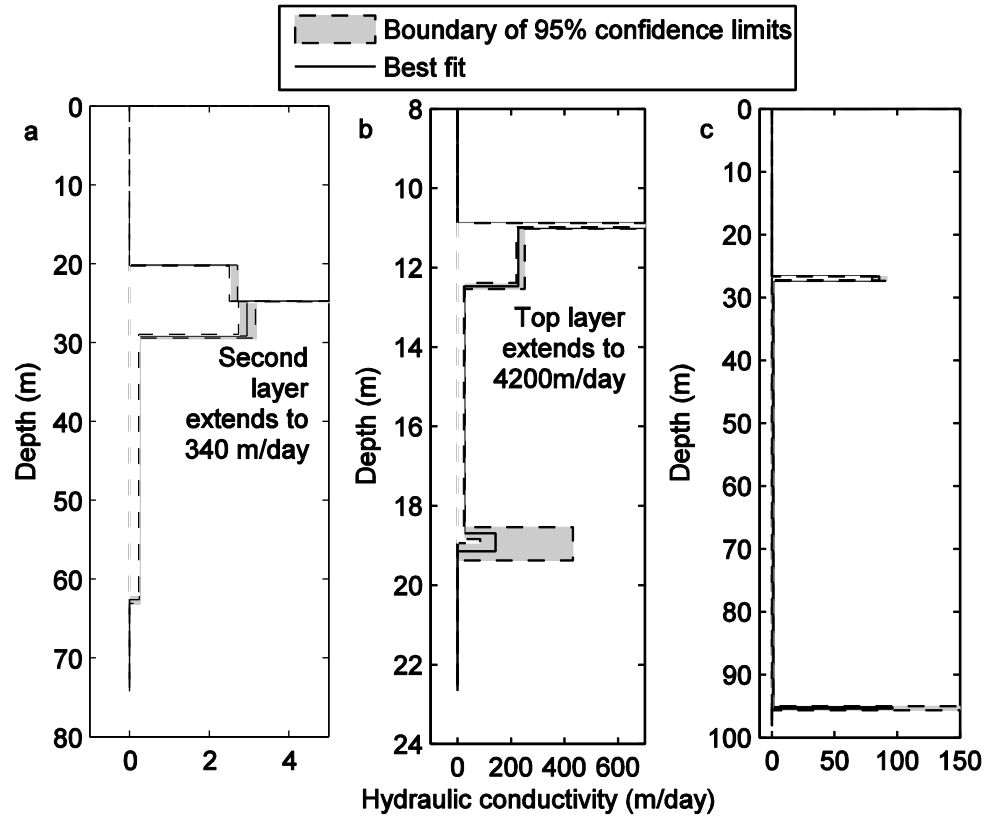
**Figure 5:** a) Benningholme b) North End Stream and c) Carnaby flow logs smoothed using a Savitzky-Golay filter



**Figure 6:** Diagram to illustrate the principle of the notch function



**Figure 7** Flow plot from (a) Benningholme and (b) North End Stream showing model fits  $n=3$ ,  $n=4$  and  $n=5$  and c) Carnaby showing model fits  $n=1$  and  $n=3$ .



**Figure 8:** Hydraulic conductivity profile for a) Benningholme b) North End Stream and c) Carnaby calculated from model data, with the 95% confidence limits for hydraulic conductivity and depth indicated. Note the value for the middle layer at Carnaby is 0.85m/day.